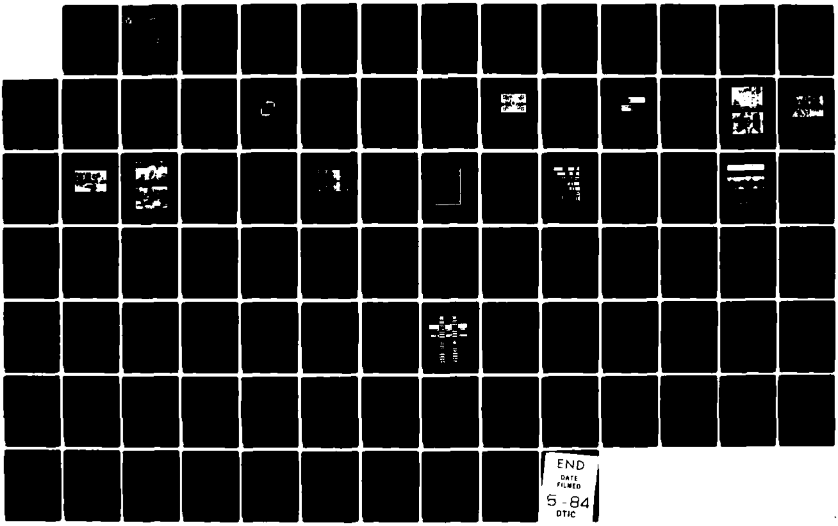
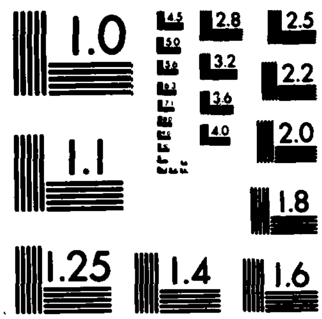
AD-A140 028 STANDARD TV IMAGES FOR QUALITY COMPARISON OF TELECONFERENCING CODECS(U) DELTA INFORMATION SYSTEMS INC JENKINTOWN PA JAN 84 NCS-TIB-84-2 DCA100-83-C-0047 1/1

UNCLASSIFIED F/G 17/2 NL




1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

(3)

NCS TIB 84-2

NATIONAL COMMUNICATIONS SYSTEM

AD A140028

TECHNICAL INFORMATION BULLETIN
84-2

STANDARD TV IMAGES
FOR QUALITY COMPARISON
OF
TELECONFERENCING CODECS

JANUARY 1984

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER NCS-TIB-84-2	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Standard TV Images for Quality Comparison of Teleconferencing Codecs		5. TYPE OF REPORT & PERIOD COVERED Final Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)		8. CONTRACT OR GRANT NUMBER(s) DCA100-83-C-0047
9. PERFORMING ORGANIZATION NAME AND ADDRESS Delta Information Systems 310 Cottoman Street Jenkintown, PA 19046		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS National Communications System Attn: NCS-TS Washington, D.C. 20305		12. REPORT DATE January 1984
		13. NUMBER OF PAGES 89
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Video, TV, Images, Teleconferencing, Still Frame, Digital Coding, Codec, Motion, Test Chart

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The purpose of this study was to identify and recommend standard TV images as the baseline for comparison of teleconferencing codecs. A series of single frame images, typical of teleconferencing applications is recommended for use as standard images for eveluating both freeze frame and motion codes. The recommended images, to be used to measure codec performance both qualitatively and quantitatively, contains test charts, people scenes, text, and color visuals.

NCS TECHNICAL INFORMATION BULLETIN 84-2

STANDARD TV IMAGES
FOR QUALITY COMPARISON OF TELECONFERENCING CODECS

JANUARY 1984

PROJECT OFFICER

Dennis Bodson

DENNIS BODSON
Senior Electronics Engineer
Office of NCS Technology
and Standards

APPROVED FOR PUBLICATION:

Marshall L. Cain

MARSHALL L. CAIN
Assistant Manager
Office of NCS Technology
and Standards

FOREWORD

Among the responsibilities assigned to the Office of the Manager, National Communications System, is the management of the Federal Telecommunication Standards Program. Under this program, the NCS, with the assistance of the Federal Telecommunication Standards Committee identifies, develops, and coordinates proposed Federal Standards which either contribute to the interoperability of functionally similar Federal telecommunication systems or to the achievement of a compatible and efficient interface between computer and telecommunication systems. In developing and coordinating these standards, a considerable amount of effort is expended in initiating and pursuing joint standards development efforts with appropriate technical committees of the Electronic Industries Association, the American National Standards Institute, the International Organization for Standardization, and the International Telegraph and Telephone Consultative Committee of the International Telecommunication Union. This Technical Information Bulletin presents an overview of an effort which is contributing to the development of compatible Federal, national, and international standards in the area of video teleconferencing standards. It has been prepared to inform interested Federal activities of the progress of these efforts. Any comments, inputs or statements of requirements which could assist in the advancement of this work are welcome and should be addressed to:

Office of the Manager
National Communications System
ATTN: NCS-TS
Washington, DC 20305
(202) 692-2124

TSI 84-02

STANDARD TV

IMAGES

FOR QUALITY COMPARISON

OF TELECONFERENCING CODECS

January 1984

Submitted to:

NATIONAL COMMUNICATIONS SYSTEM
OFFICE OF TECHNOLOGY AND STANDARDS
ARLINGTON, VA 22204

Contracting Agency:

DEFENSE COMMUNICATIONS AGENCY

Contract Number
DCA100-83-C-0047

Submitted by:

DELTA INFORMATION SYSTEMS, INC.
310 Cottman Street
Jenkintown, Pa. 19046

TABLE OF CONTENTS

List of Figures

1.0 Introduction

This report summarizes the work on Task 5.0, "Standard TV Images", performed by Delta Information Systems, Inc. for the Office of Technology and Standards of the National Communications System, an organization of the U.S. Government, under contract number DCA100-83-C-0047. The Office of Technology and Standards, headed by National Communications System Assistant Manager Marshall L. Cain, is responsible for the management of the Federal Telecommunications Standards Program, which develops telecommunications standards whose use is mandatory by all Federal agencies.

A series of single frame images, typical of teleconferencing applications, is recommended for use as standard images for evaluating both freeze frame and motion codecs. The recommended images, to be used to measure codec performance both qualitatively and quantitatively, contain test charts, people scenes, text and color visuals. This work is described in Section 2.

The method of converting the recommended imagery to digital format is described in Section 3. A discussion of various digital recording media is contained in Section 4 where a digital tape is recommended. Section 5 recommends that a 1" type C playback video tape player be used to record and play the analog test sequences.

In Section 7.0, performance criteria and parameters which are affected by the codec motion response is discussed. Techniques are described for measuring motion response and motion sequences are recommended

Methods of measuring codec performance by qualitative and quantitative testing are described. Evaluation procedures, comparative testing connections and measurement rating factors are recommended in Section 9.0. Recommendations for further work in developing standard images for High Definition Television (HDTV) are described in Section 11.

2.0 Development of Standard Still Frame Test Images

2.1 Purpose

The purpose of the standard still frame test images is to provide a consistent source of test signals which stress the various parameters of still frame image transmission systems designed for teleconferencing. This permits evaluating the performance of these systems on an absolute or comparative basis. It eliminates the variability in test signals from test to test when they are generated "live" for each test. These variabilities can be substantial; for example, in shading, resolution due to optical and electrical focus and/or astigmatism, signal-to-noise ratio due to scene illumination or camera adjustment, geometric anomalies due to raster size and non-linearities, etc.

It is, of course, desireable to minimize any distortions of the type described. This can be done in the generation of a "standard" test tape because the best TV camera can be procured and great pains taken since this is a one-time effort. Once these signals have been optimized they are converted to digital signals and recorded on archival tapes. Since the tapes are digital, the quality of the recovered signal is consistent on each replay. The analog tapes subsequently generated from the digital signals will all produce identical signals of the highest quality so that all tests performed are referenced to identical signals. Thus the tests are truly meaningful.

The test images are designed to stress the parameters of the systems being evaluated. Ideally the content of the test signals will range from that which is easily conveyed by the system, through that which is at the limit of the system capability, to that which is beyond the capability of all but the finest of systems or may not be within present capabilities but may be handled by future systems.

2.2 Parameters to be Evaluated

The parameters of the image transmission system which must be evaluated to determine the performance of the system fall into two categories; those which are usually evaluated quantitatively and those which can be evaluated qualitatively. These categories are not clear cut and some of the parameters may fall into both categories depending on the structure of the test. The categories will be used here only as an aid in tabulating all of the parameters to be evaluated as follows.

The definitions of these parameters are well-known and understood in the industry and therefore will not be reiterated here.

2.3 Quantitative Parameters

The following is a tabulation of parameters which can be evaluated quantitatively.

- Gray Scale
- Resolution
  - Horizontal
  - Vertical
  - Slant Edges
- Quantization
- Color
  - Saturation
  - Purity
  - Hue Accuracy
  - Crosstalk
  - Effects of Color on Luminance
  - Effects of Luminescence on Color
- Signal-to-Noise Ratio
- Temporal Effects
  - Frame Grab
  - Retention of Preceding Images

2.4 Qualitative Parameters

The following is a tabulation of parameters which can best be evaluated qualitatively.

Flesh Tone of Subjects

Recognizability of Subjects

Effects of Scene on Appearance of Subjects

Clarity of Print, Graphs, and Maps

Ability to Specify Single Colors

Ability to Determine Colors in Multi-Color Scenes

S/N or S/Artifact Performance

2.5 Test Images Selected

The test images are selected to evaluate the parameters listed above to determine the performance capability of the equipment being tested. The test images selected generally consist of scenes or other visuals typical, or very similar, to those which will be transmitted by the equipment being tested. These include primarily human subjects in a field of view typical of teleconferencing requirements. For example, a) a close up of a single person, b) medium shot of two people, and c) wide angle shot with two people. The background contains artifacts of a type and size encountered in teleconferencing. The original slide material was very carefully designed and produced, in both content and technical characteristics such as color, density, and density range. Therefore the evaluation results should be indicative of equipment performance in its intended application.

2.5.1 800 Line Test Chart

The 800 line test chart selected, Figure 2-1, is the EIA Resolution Chart 1956 (formerly RETMA). This chart is probably the single most meaningful test image. It was designed to provide a standard reference for measuring the resolution of television systems and as an aid in testing for streaking, ringing, interlace, shading, scanning linearity, aspect ratio, and gray scale reproduction. It is also useful in detecting aliasing.

The center horizontal and vertical wedges are composed of four black lines separated by three equal width white lines. The numbers printed alongside the wedges correspond to the total number of lines (black and white) at that pitch which can be placed adjacent to one another in the height of the chart. Since the aspect ratio is 4:3, 4/3 times this number is the number of lines which can be placed in the width of the chart. The frequencies produced in the television signal by the vertical lines are 1.25 megaHartz per 100 lines as printed next to the wedge or 80 lines per megaHertz. The following table shows some representative examples.

Line Number of Vertical Wedge	Fundamental Video Frequency
200	2.5 MHZ
240	3.0
280	3.5
320	4.0
400	5.0
480	6.0
560	7.0
640	8.0
720	9.0
800	10.0

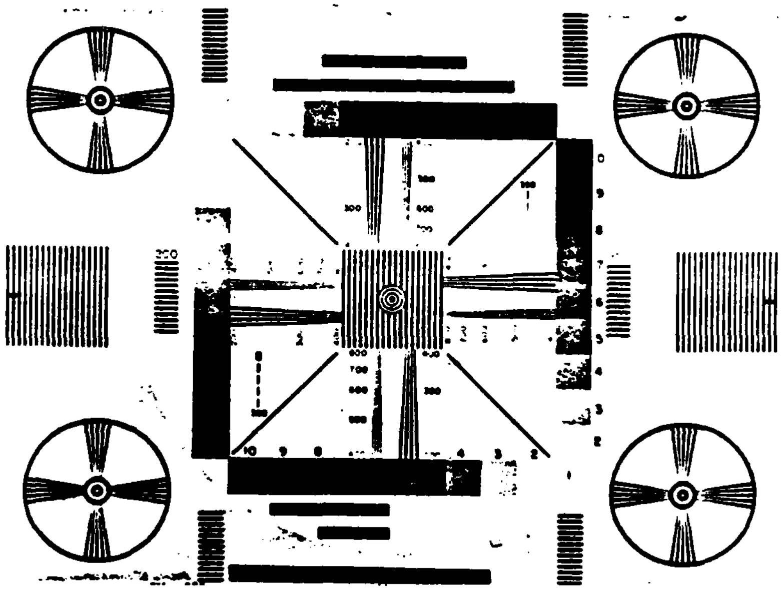


FIGURE 2-1 800 LINE TEST CHART

Obviously the digitization process required to generate the digital archival tape and the subsequent user analog test tape will truncate the wedges on a theoretical basis at about 450 lines. It is recommended that a direct analog signal produced electronically be recorded on the analog user tape to be used for resolution and aliasing measurements if the requirement for higher precision arises. Since the signal is electronically generated, its quality will be consistently reproducable.

The ten step gray scales cover a contrast range of approximately 30:1. Step #1 is clear; that is, it is the same density as the center circle and is used as the "white" reference. The remaining nine steps are calibrated with a density difference of 0.16 between steps. When the test chart is in the form of a chart rather than a transparency, paste-on calibrated gray scales are used which have the following specifications.

Gray Scale Number	Nominal Reflectance	Nominal Reflection Density
1 (circle)	Greater than 60%	Less than 0.22
2	60	0.22
3	41.7	0.38
4	28.2	0.55
5	19.5	0.71
6	13.5	0.87
7	9.3	1.03
8	6.3	1.20
9	4.4	1.36
10	3.0	1.52

Since the camera tube response is not linear, gamma correction circuits can be adjusted at operator discretion, and lighting can vary widely, meaningful tests can be made only with the controlled, standardized test tape providing identical signals for each test.

Quantization may be checked by visual inspection of the picture monitor to determine if the background on the steps within the gray scales are constant in gray level or vary smoothly with no indication of contouring.

Low frequency response is indicated by observing the horizontal black bars for trailing streaking effects.

The four diagonal black lines inside the square formed by the gray scales are useful for checking interlace. A jagged line indicates pairing of the interlaced lines. This further checks the relationship and/or presence of the odd and even field in the read-out of single frame pictures from the internal memory of the equipment being evaluated.

Ringing due to a variety of causes in the analog or the digital circuitry or cross coupling is manifested in the single line widths located in the upper right and lower left hand portions of the central area. The lines in the upper right have widths corresponding to 350 to 550 lines. Those in the lower left have widths from 100 to 300 lines.

2.5.2 Standard Test Chart

The standard test chart is included primarily as a convenience in determining resolution. It contains four resolution wedges substantially larger in area than those of the newer 800 line test chart. The line numbers vary from 200 to 500 in each of these wedges. Four smaller additional sets of resolution wedges are located in the corners of the test chart. Their line numbers vary from 200 to 400. The large uniform background area provides an excellent test for contouring. The central large and four smaller corner circles can provide information on sampling uniformity and interlace. This test chart contains only three density levels. The standard test chart is shown in Figure 2-2.

2.5.3 Gray Scale Test Chart

This test chart consists of two calibrated gray scales considerably larger in size than those on the 800 line test chart. The gray scales are arranged horizontally, one above the other, with the gradation in density in opposite directions on a neutral background. The gradation steps are as defined in Section 2.5.1. This arrangement assists in detecting variations in shading across the image area.

This test chart verifies the accuracy in quantization and is ideal for checking contouring. It can also be used for S/N determination.

The gray scale test chart can also be generated electronically. In this case it will not have the slight shading introduced by a TV camera and as such would not represent an actual signal encountered in practical operation. The electronically generated

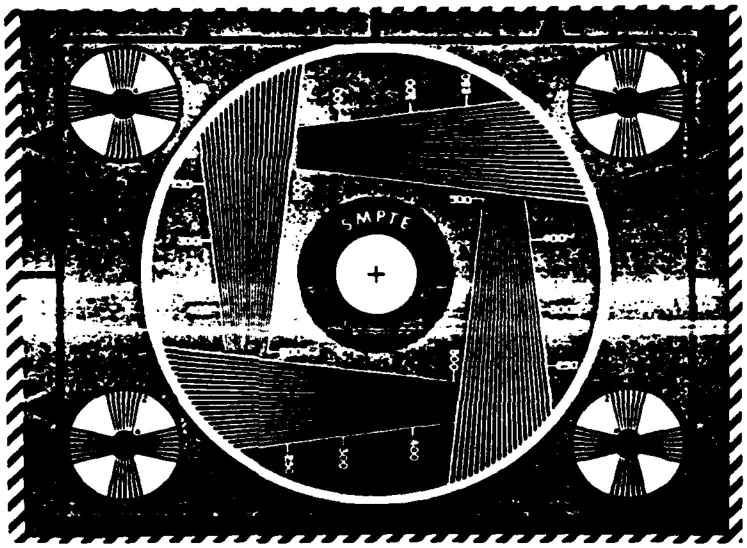


FIGURE 2-2 STANDARD TEST CHART

gray scale might be very useful for system design, debug, and alignment. The adviseability of including this signal will be considered. See Figure 2-3.

2.5.4 Single Person, Close-Up

Two test images are included of a single person in a close-up view, one male and one female. These images are indicative of a single speaker in a teleconferencing presentation. The person's face is at the plane of optimum focus. The field of view at this point is about 2 feet wide. The luminance range as measured with a spot luminance meter is about 25:1. Using reference white as 100% relative luminance, the female face is at 30-35% and the male face at 20-30%. The lowest measureable dark area is at 4% relative luminance. These slides are SMPTE color television Subjective Reference Slides Numbers 4 and 5. The slide of the female face has been used for television system evaluation for about 30 years and therefore has the added feature that its transmission can be very accurately evaluated by anyone experienced in video systems.

These test images are excellent for the qualitative evaluation of most of the parameters involved in image transmission for teleconferencing. Close-up scrutiny of highly detailed areas such as the hair, eyes, and teeth, reveal the practical effects of system resolution and quantizing precision. The gradually changing tones of the face, the lady's scarf, and the man's jacket, as

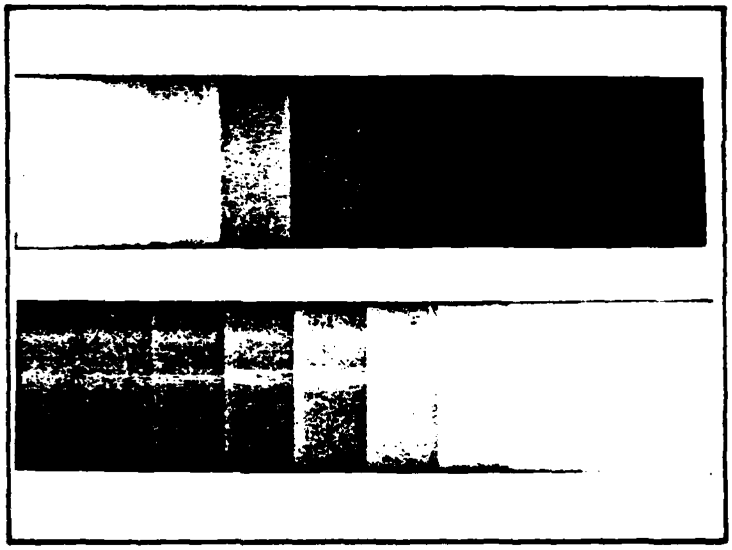


FIGURE 2-3 GRAY SCALE CHART

well as the out-of-focus background are indicative of quantization accuracy and contouring. Flesh tones of the face provide excellent evaluation of low saturation color while the lady's flowers and scarf represent comparatively high saturation colors. Thus many aspects of practical color performance can be easily evaluated. These test images are shown in Figure 2-4.

2.5.5 Two People, Medium Shot

This test image depicts two people standing in a field of view of about 4 feet by 3 feet. This field of view and the relative sizes of the faces compared to the total field of view is about the same as would be encountered in a teleconferencing transmission of two people seated at a desk or, as in the image, standing during a presentation. Since all of the features of this test image are relatively smaller than in the single person test image, the various technical parameters are stressed by different areas of the scene providing an independent evaluation of each in a practical application.

The technical parameters of the actual test image, luminance and luminance range, are the same as the single person test image. This test image is the SMPTE Color Television Subjective Reference Slide Number 3. It is shown in Figure 2-5.

2.5.6 Two People, Wide Angle

This test image shows two people in a room. The field of view at the plane in which the people are standing is about 10 feet

FIGURE 2-4 SINGLE PERSON TEST IMAGES

FIGURE 2-5 MEDIUM SHOT OF TWO PEOPLE TEST IMAGE

by 15 feet. This corresponds to a teleconferencing scene in which two people are moving about in an area containing objects which are subjects of the teleconference; eg, product demonstration, etc. The scene contains areas of gradually changing luminance and chrominance of low saturation as well as abruptly changing areas of high luminance and chrominance contrast. The faces of the subjects are relatively small. Objects are located in the foreground and background, as well as in the focal plane. Luminance ranges are 25:1 as previously described. This scene is therefore excellent for testing all of the technical parameters listed qualitatively.

This test image is SMPTE Color Television Reference Slide Number 2. It is shown in Figure 2-6.

2.5.7 Chroma Test Images

The chroma test image shown in Figure 2-7 is of a young lady working in a kitchen. Wardrobe and set colors used in this scene provide good color contrast but relatively low brightness contrast. Therefore chromatic capabilities and the effects of chroma on luminance can be evaluated. This is SMPTE test Slide 15.

A second test image for this purpose, shown in Figure 2-7, is of a girl in a green dress against a neutral background. The large area background should appear constant in luminance and color thus further testing the effect of chroma cross-talk. This is SMPTE test Slide 14.

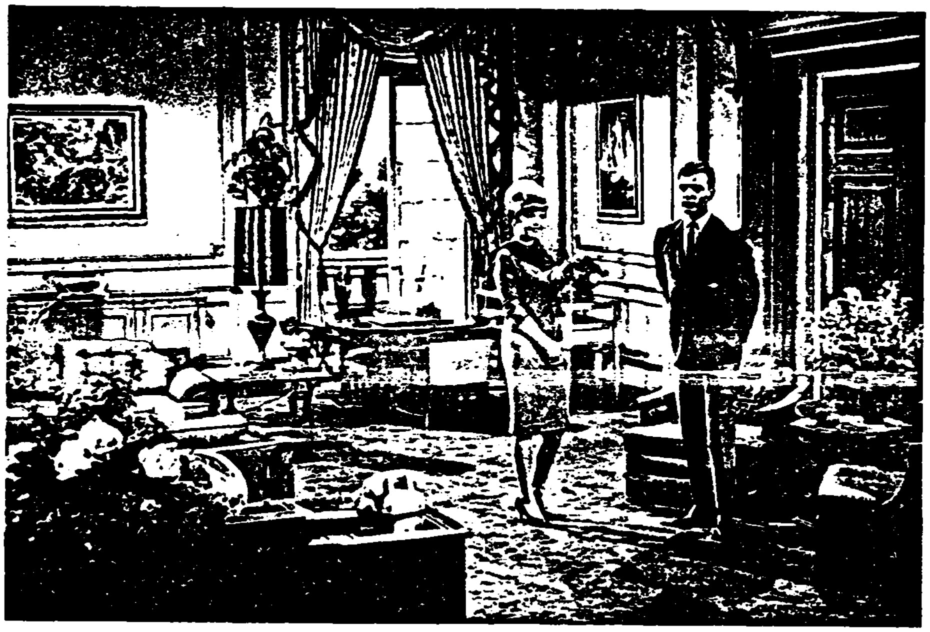


FIGURE 2-6 WIDE SHOT OF TWO PEOPLE TEST IMAGE

FIGURE 2-7 CHROMA TEST IMAGES

2.5.8 Black-White Text

The black-white text test image consists of alpha-numeric groups of various fonts and sizes as shown in Figure 2-8. By means of this test slide the ability of the luminance sampling concept and hardware capability of the system under evaluation to reproduce text can be thoroughly evaluated. The fonts chosen are those most commonly used in television titlers in broadcast and teleconferencing applications. The character sizes chosen are 30, 22.5, 14.5, and 11 scan lines per character. Larger characters, although sometimes used, do not test the system. The reasonable limit for character size in an NTSC system is 10 scan lines per character. Therefore smaller characters are not included.

2.5.9 Color Text

The color text test image consists of alpha-numeric groups of various fonts and sizes arranged against different color backgrounds as shown in Figure 2-9. By means of this test slide the ability of the luminance and chroma sampling concept and hardware capability of the system under evaluation to reproduce text can be thoroughly evaluated.

2.5.10 Black-White Graphs

These test images consist of bar and line graphs. The bars in the bar graph consist of various types of crosshatch for segment contrast and the lines are at various angles. The lines in the

FIGURE 2-8 BLACK-WHITE TEXT TEST IMAGE

FIGURE 2-9 COLOR TEXT TEST IMAGE

line graph go through a wide variety of slopes alternately presenting line segments which are narrow horizontally and then vertically with rapidly changing slopes between. These slides thoroughly test the luminance quantizing precision, sampling standard and interlace of the system under evaluation. They are shown in Figures 2-10 and 2-11.

2.5.11 Color Graphs

These test images consist of bar and line graphs identical to the black-white graphs. The bars in this case have no cross-hatching. It is replaced by solid color segments. The lines in the line graphs are supplemented by color of adequate width to permit proper performance in an NTSC system. These slides test the luminance and chrominance quantizing and sampling structure of the system under evaluation. They are shown in Figures 2-12 and 2-13.

2.5.12 Color Bar Chart

This test image is equal in importance to the resolution test image for evaluating the performance of image transmission systems. It is an essential test signal in the design, debug, and test of the hardware. For this purpose it is electronically generated by a piece of test equipment which is generally available.

However, because of its importance it is proposed that it be included as a test image on the user tape. The Color Bar Chart

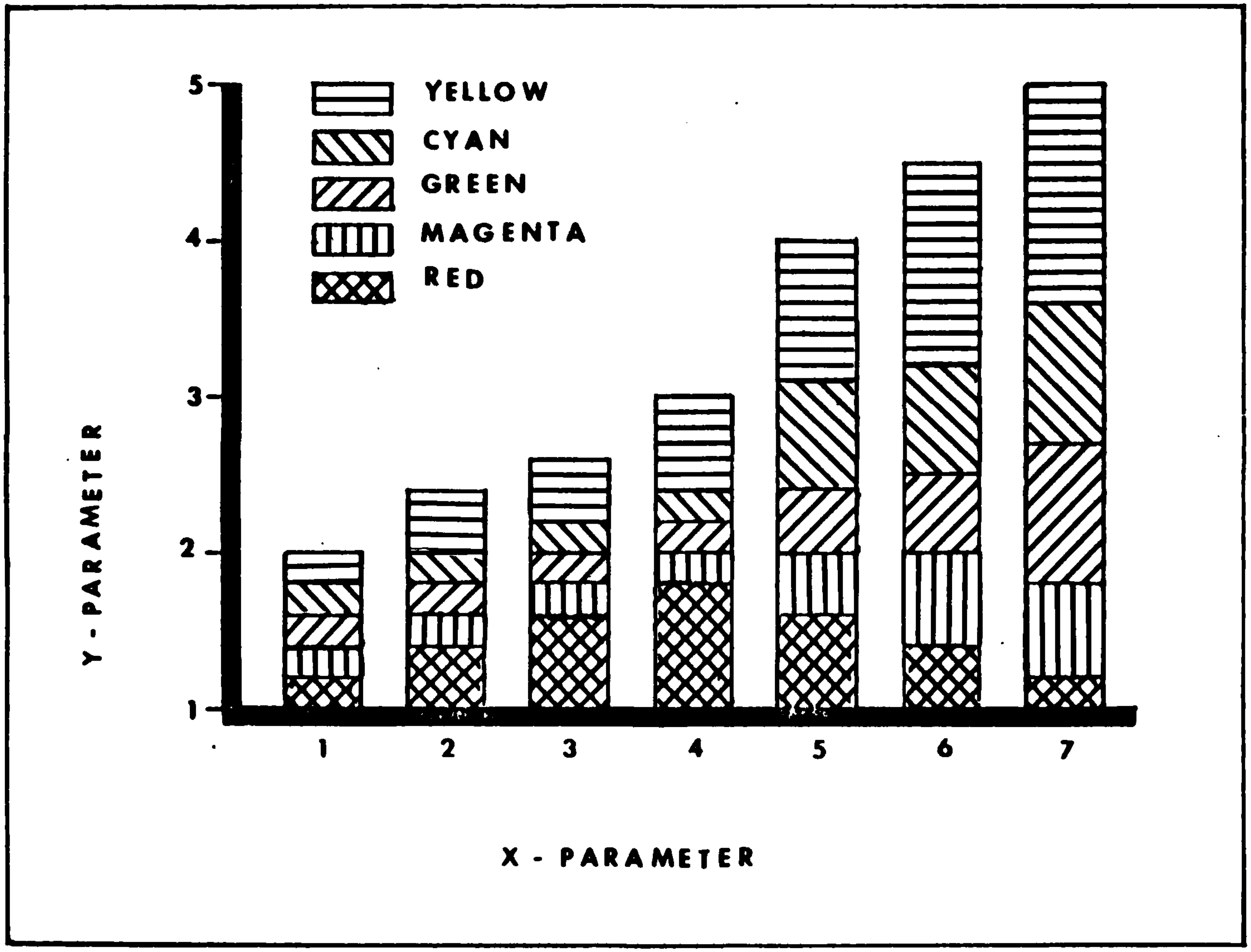


FIGURE 2-10 BLACK-WHITE BAR GRAPH TEST IMAGE

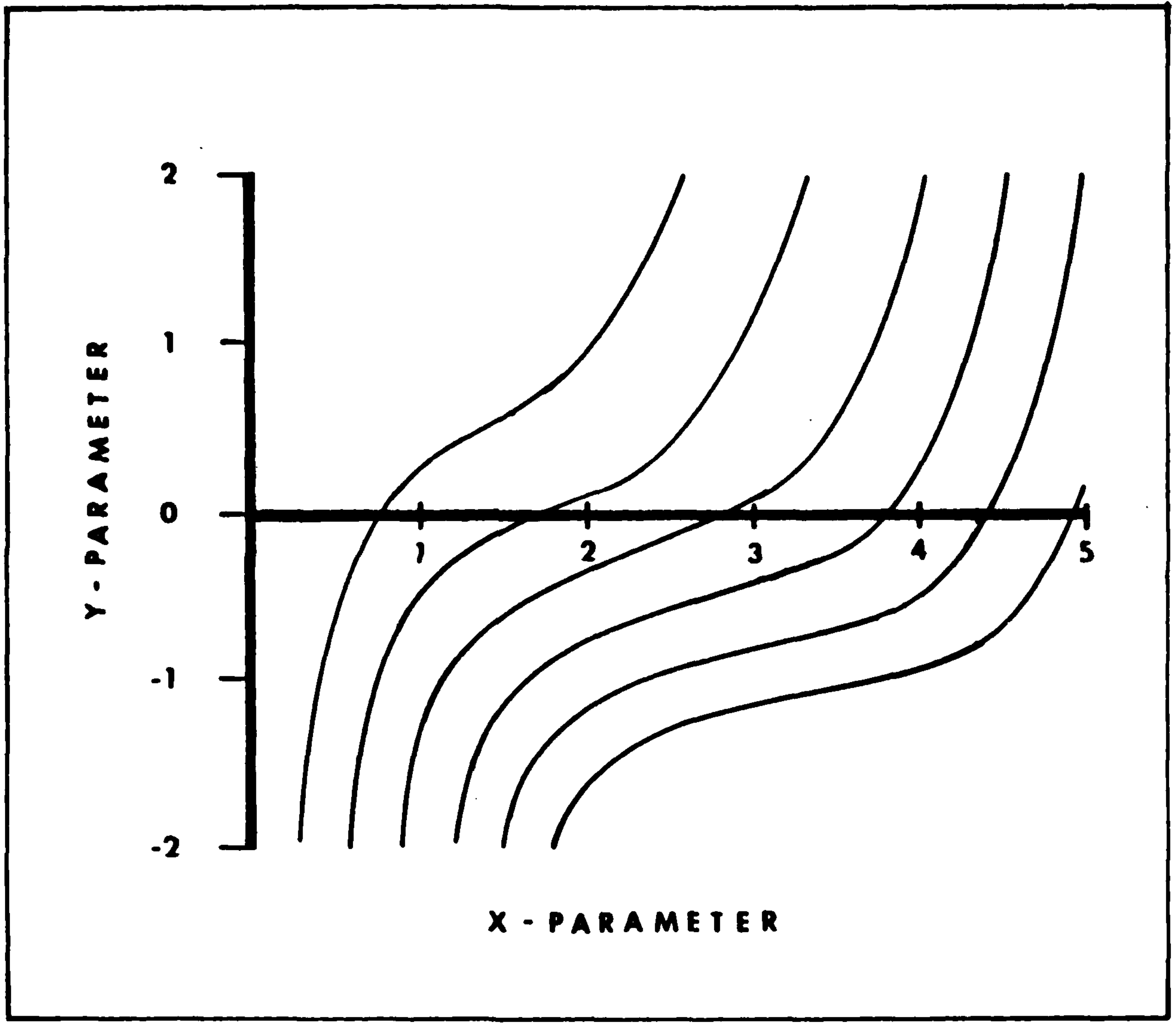


FIGURE 2-11 BLACK-WHITE LINE GRAPH TEST IMAGE

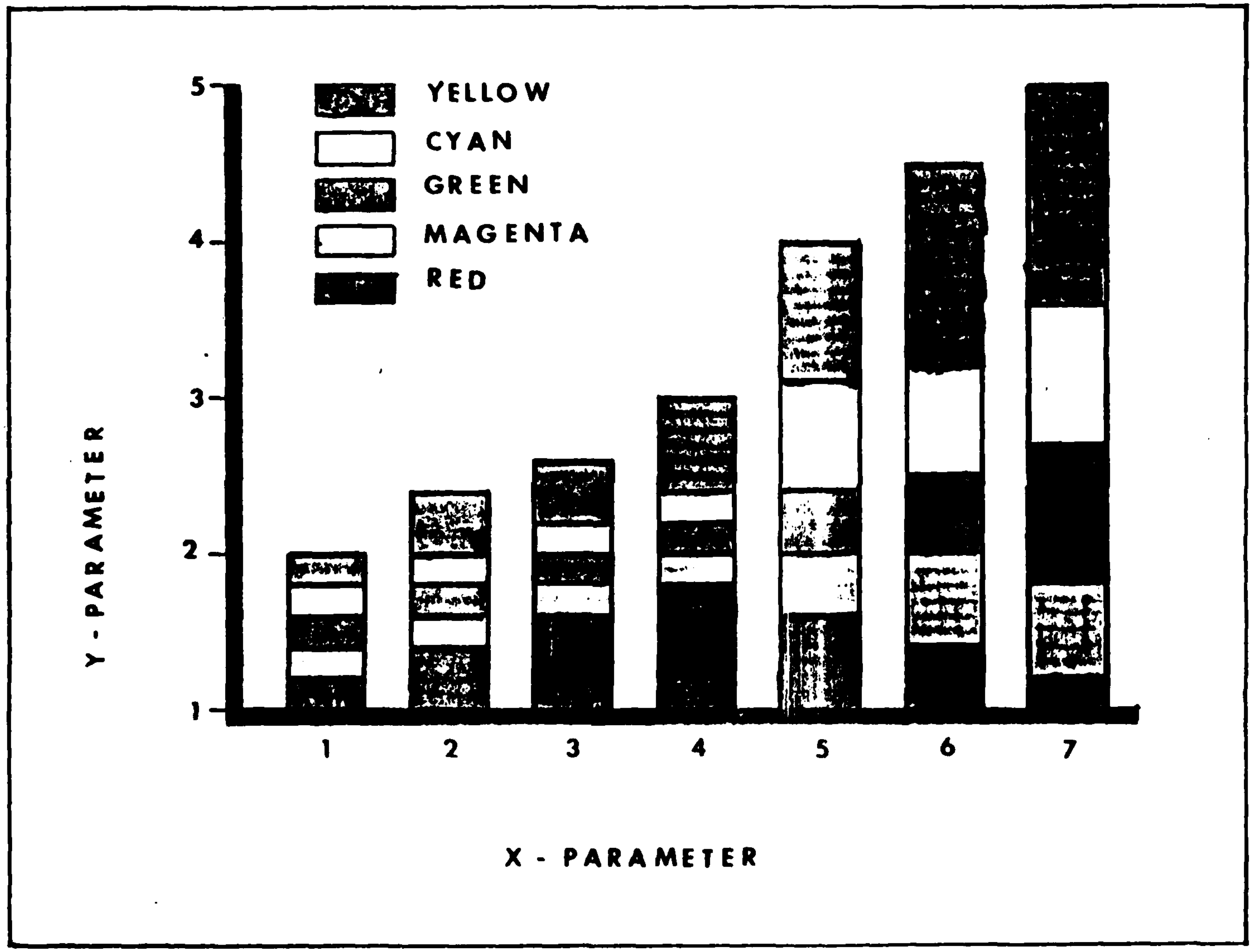


FIGURE 2-12 COLOR BAR GRAPH TEST IMAGE

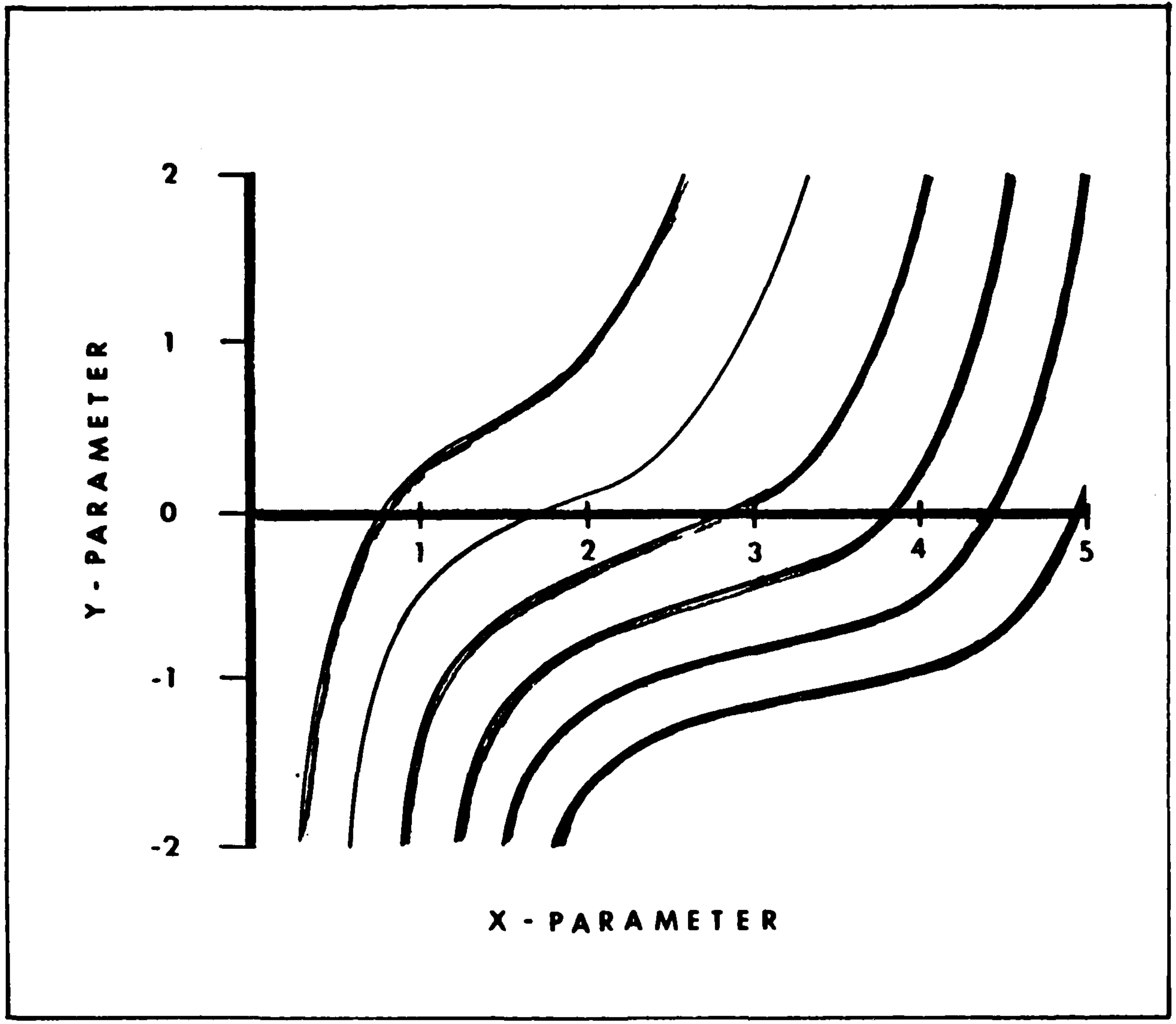


FIGURE 2-13 COLOR LINE GRAPH TEST IMAGE

is shown in Figure 2-14.

2.5.13 Aliasing Test Image

The anti-aliasing filter is a very critical element of the image transmission system. By the very nature of its principal function, preventing aliasing, it controls the entire video pass-band as to phase and frequency response. These effects can be detected with the 800 line test chart. The adequacy of aliasing prevention may require a separate test image as described below. A test image of alternating vertical black and white bars continuously increasing in pitch from left to right can be used to easily detect aliasing or other sampling anomalies. It is essentially a constant amplitude swept frequency signal coherent with the scanning standard. As such it can be used to evaluate the phase and frequency response of the video pass band and the anti-aliasing performance as well as sampling anomalies.

This signal is similar to the familiar multi-burst signal but is continuous in frequency so that all frequency components of the video pass band can be evaluated. It is easily generated electronically and therefore repeatable. It is recommended that it be added to the user analog test tape.

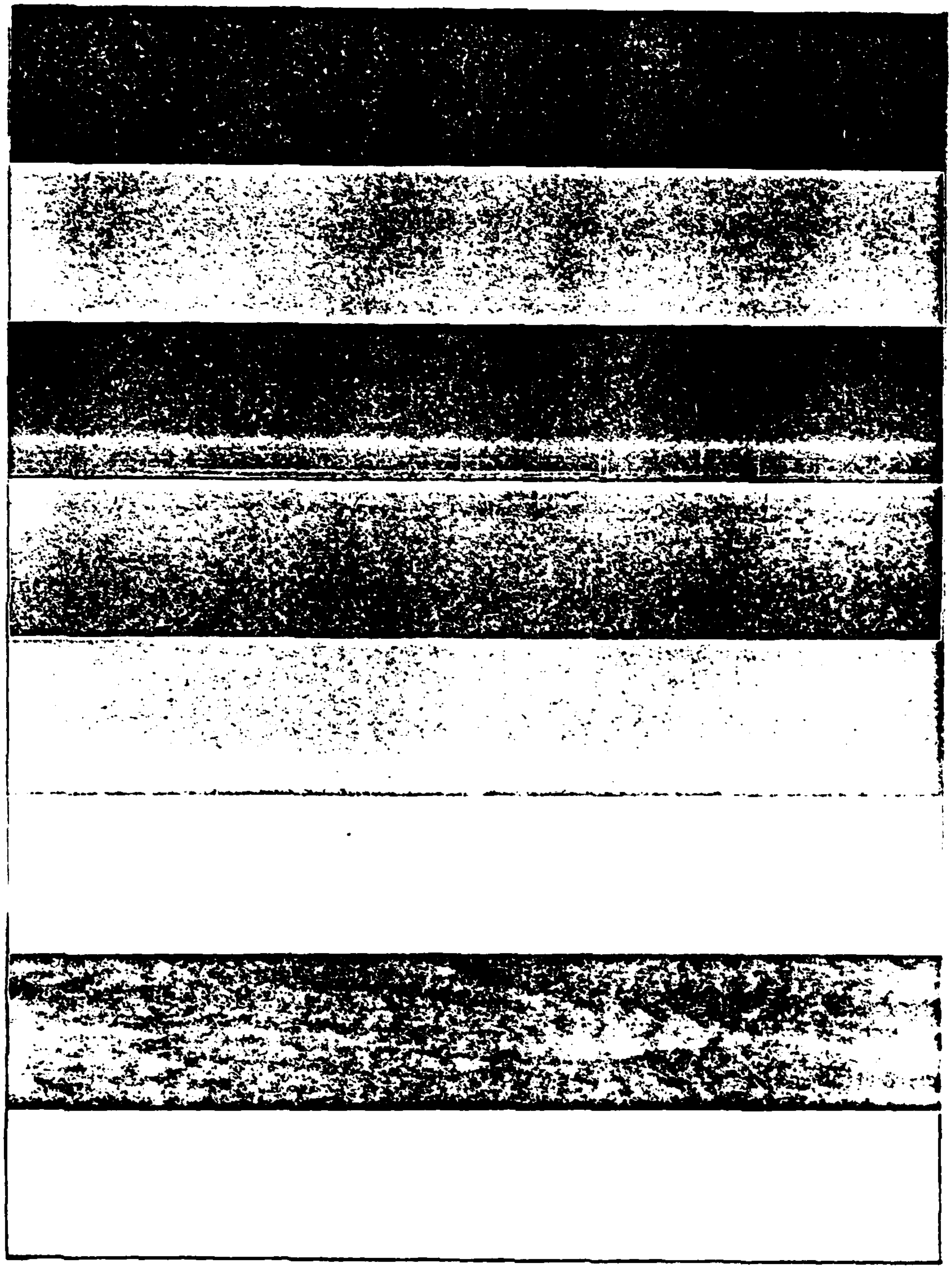


FIGURE 2-14 COLOR BAR CHART

3.0 Digital Coding Recommendation

3.1 Purpose

As previously described, it is essential to store the master set of test imagery in digital form so that the various causes of signal degradation associated with analog storage will not affect the future quality and consistency of the reproduced signal. The original video signal is analog as is the final version of the test signal to be used as the codec test signal. Therefore the video signal must be converted to digital form for storage and subsequently back to analog form for future use. The format of the digital signal must fulfill two basic requirements.

1. It must be compatible with the digital storage medium selected.

2. It must be compatible with the video signal; that is, it must neither degrade the quality of, nor inject artifacts into, the video signal.

In summary then, the purpose of the digital coding is to permit converting an analog video signal of broadcast quality into digital format without degrading the image so that it may be stored and recalled many times for conversion back to analog form, producing signals of consistent quality.

3.2 Types of Digital Coding

Digital coding techniques may be categorized in several ways. The most important of these are

a) compression, and

b) the form of the input signal.

3.2.1 Compression

Compression is a term used to categorize a digital coding technique which produces fewer data bits to specify an image than required by a uncompressed coding scheme (e.g. PCM). The uncompressed coding scheme is one which can be used to digitally code the analog video signal of any image of a given quality with either no degradation or with only specific agreed upon degradation. The purpose of compression, of course, is to define the image using fewer data bits for such purposes as minimizing transmission time or storage volume. One figure of merit of a compression system is compression ratio defined as the ratio of the number of uncompressed data bits to the number of compressed data bits for a given image. The types of compression techniques available is very extensive, each with its own advantages and disadvantages It is not necessary to consider them as shown below.

Having defined the categories of compressed and uncompressed digital coding techniques, it will now be shown that only uncompressed coding should be considered for these reasons.

1. It produces the highest and most consistent quality results.
2. The prime advantage of compression techniques is fewer data bits per image. This is not important to the archival storage

since time and volume do not appear to be a problem.

3. Coding techniques generally capitalize on the redundant, statistical, or temporal characteristics of images in general to reduce the number of bits defining an image. The coding techniques are generally optimized for certain types of image characteristics and as a result may produce degradation in other characteristics often in relationship to the compression ratio achieved.

Therefore since storage volume is not a problem and consistent quality is essential, compressed coding techniques have been rejected in favor of an uncompressed digital coding technique.

3.2.2 R-G-B, NTSC, Component Coding

The second way in which digital coding techniques may be categorized is by the form of the analog signal on which they operate. Principal among these are:

1. RGB
2. NTSC
3. Luminance and chrominance components.

Probably the ideal input signal would be wide band red, green, and blue video signals. This may be considered the parent video signal format from which NTSC and the component form of video signal are generated, each with somewhat lower quality through very well designed and thoroughly specified acceptable degradations.

There are two reasons why the R, G, B input video signal is not required. First, it will rarely be the input for which the codec under evaluation was designed. Except for customized units, the codecs are designed for NTSC (or occasionally component) input video signals. Second, wideband R, G, B signals are generally not available as an accessable output of a television camera each time new evaluation images are to be generated.

In some cases the components of the video signal (Y, R-Y, and B-Y) are available. In almost all cases the NTSC form is available for input to the digital encoding system. It has been shown that better control of the digitally encoded video signal is possible when operating on the component form of the video signal. It is generally agreed that sampling and coding of the NTSC signal directly requires locking the sampling rate to the color subcarrier. However, high quality conversion of the NTSC signal into component form can be incorporated into the digital encoder with little effort so that consistent quality component signals are available as the baseline signal for the actual encoding equipment.

Thus, it is concluded that the signal be stored as a digitization of the component form of the signal as decoded from an NTSC video source.

3.2.3 Coding Parameters

The principal coding parameters to be defined are sampling rate, filter characteristics, quantizing precision, and timing relationships.

Prior to discussing each of these parameters it may be well to say that a tremendous amount of effort was expended initially by industry and more recently by the EBU and the SMPTE in evaluating these requirements. In what follows it will be shown that the EBU and the tentative SMPTE standards are extremely well designed and therefore close adherence to the specified parameters is recommended.

3.2.3.1 Sampling Rate

The minimum required sampling rate is determined by the horizontal resolution which the digital system must support. Vertical resolution is, of course, determined by the television system line structure which, in effect, is also a sampling system. In a television system employing about 480 active lines in each frame, the maximum vertical resolution is 480 picture elements (PELS). This is modified by the statistical Kell factor whose value is between 0.7 and 0.9. Using 0.7 for the Kell factor, the effective vertical resolution is 336 PEL. The digital coding system has little effect on vertical resolution as long as it includes all the active lines of the frame in the conversion process.

The NTSC system was designed to provide equal horizontal and vertical resolution. In the NTSC system the active line interval is about 51.5 microseconds and the video bandwidth is

4.2 megaHertz. Thus a maximum of 216 cycles can occur within the active line interval to support 432 PELS. Since the height of the image is 3/4 of the width, 3/4 x 432 or 324 PELS can occur in an image width equal to its height. So the goal of equal horizontal and vertical resolution are met. To support 432 PELS the digital system must sample the video signal at least 432 times within the active line interval. This criteria will digitize continuing segments of equally spaced white and black bars as found in resolution test charts or multi-burst signals in the rare case where samples and bars are in phase. However, to reproduce randomly located, single, isolated bars and edges with good fidelity requires oversampling (similar to the Kell Factor). Experience has shown that sampling rates above 700 per active line can support this requirement. The present SMPTE recommendations and the adopted EBU standards are 720 samples per active line interval for luminance.

The chrominance resolution provided by an NTSC signal is substantially lower. The maximum bandwidth is 1.2 megaHertz which will support 124 PEL per active line. Both of the above organizations have elected to sample the chrominance signals, R-Y and B-Y, 360 times per active line interval. This permits locking the chrominance sampling to the luminance sampling by a factor of 2:1.

It is recommended that the digital coding standard used employ sampling rates which provide equivalent performance to those discussed above.

3.2.3.2 Filter Characteristics

A filter is required in the digital coding system to remove components within the video signal which could cause aliasing in the conversion process. These are video signal components with frequencies above the sampling rate. The sampling process causes these components to appear as artifacts in the reconstructed image. A filter is therefore employed as an essnetial component of the digital coding process to remove all video signal components which could produce aliasing. The filter should provide full pass band up to 4.2 megaHertz and roll off as sharply as possible to provide a high degree of attenuation at, and above, the sampling frequency.* The sampling frequency required to provide in the order of 720 samples in an active line interval is about 13.5 megaHertz. Digitization of the chrominance signals requires similar filters. In this case the sampling rate is 6.75 megaHertz. Practical filters are available for both requirements.

*The filter characteristics within the pass band are critical particularly with regard to phase and attenuation.

3.2.3.3 Quantizing Precision

Quantizing precision is defined as the number of levels which the digital codec can differentiate and define within the video signal. It is determined in two ways; qualitatively, and quantitatively. Qualitatively the effect of quantizing precision appears as contouring of similar gray levels in the reproduced image. It is generally agreed that contouring is visible when 64 levels are defined and occassionally minimally visible at 128 levels. It is agreed that with 256 levels contouring is not visible in a camera generated television signal. Using a higher precision, 512 is not considered necessary for this type of video signal. The 256 levels can be defined by 8 digital bits. Based on the preceding, 8 bit precision is recommended for both the luminance and chrominance signals.

3.2.3.4 Sample to Video Timing Relationship

The preceding sections have rationalized that the preferred method of digital coding is to operate on the Y, R-Y, B-Y components of the video signal. The relationship of the sampling field to the image is less critical for this method as compared to direct sampling of the encoded NTSC signal. An accepted configuration of the sampling field for the 4:2:2 sampling format (as defined by the EBU and SMPTE digital video standards and shown below as Standard 1) is that each active line of luminance is sampled 720 times. The corresponding line of R-Y and B-Y are each sampled 360 times. The only relationship specified is that the chrominance samples coincide with the odd luminance samples; that is, the 1st, 3rd, 5th, up to the 719th.

3.3 Selected Digital Coding Recommendation

The parameters derived in the preceding sections are used as the guidelines to the recommendation defined below. Two recommendations very similar to each other are shown in the tabulation because hardware exists for both and both provide comparable quality.

	Recommendation 1	Recommendation 2
Coded Signals	Y, R-Y, B-Y	Y, R-Y, B-Y
Number of Samples per total line		
luminance	858	922
chrominance	429	230
Sampling frequency		
luminance	13.5 MHz	14.3 MHz
chrominance	6.75 MHz	3.58 MHz
Quantization Precision	8-bits	8-bits
Quantization	Uniform	Uniform
Number of Samples per active line		
luminance	720	768
chrominance (each)	360	192

4.0 Digital Recording Medium

4.1 Purpose

The purpose of the digital recording medium is to provide archival storage for the test video images which is essentially free from degradation effects of multiple replay, environment, or medium age. In addition, it is essential that second generation recordings can be made with no change in test image quality. Several digital storage mediums can provide this capability, each with its distinctive advantages and disadvantages. Principal among these are digital tape, digital disc, and solid state.

4.2 Types

4.2.1 Digital Tape

Digital tape is an excellent medium for archival storage of the test images. It is a highly developed medium which, with reasonable care will minimize any degradation due to multiple replay, environment, or medium age. Multiple copies can be made with no change in test image quality (eg. S/N ratio) using standard digital transfer equipments. The principal advantage of digital tape is that it is highly transportable. Proper selection of the recording format will permit its replay from many standard tape transport equipments. A secondary advantage is that digital tape is a low cost storage medium. The principal disadvantage is that the tape readout is comparatively slow. Furthermore, the data must be sectionalized. First, data on digital tape is not continuous but occurs in groupings to facilitate the location, recording, playback, and error checking processes. Secondly, since data occurs in descrete entries (PELS) it is sectionalized by Y, R-Y, B-Y entries. This can occur sequentially by PEL or by section, one for Y, one for R-Y, and one for B-Y. Thus the image is stored in data format rather than in image format. The result of the comparatively slow readout and the sectionalization of data with digital tape requires that the image be reassembled into image format by a subsequent device so that it can be read out at video rates and in analog form for actual use (the generation of the analog test tape).

4.2.2 Digital Disk

Digital disk storage is a very popular digital storage method. It occurs in two general forms; hard disk and soft disk (more generally referred to as floppy disks).

The hard disk can provide adequate capacity to store a full digital image on a single disk. This type of disk is not an ideal transportable module. It prefers to reside in a parent machine with the appropriate driver modules.

The floppy disk is readily transportable. It is used with many of the personal computers so that it is not as machine limited as the hard disk. It's capacity is in general inadequate to store a complete test image. As a low cost, convenient archival storage medium, it provides a second choice to the digital tape.

4.2.3 Other Devices

Principal among other types of digital storage devices is the solid state memory. This concept of digital storage is not presently practical for archival storage because the storage requires that power be applied at least at a minimal level for the duration of the storage period. Further, cost of solid state storage is prohibitive and it is not transportable in the sense of video tape.

However, a solid state memory is essential as the speed buffer and reassembly device in reconstructing the digitally stored image for subsequent use.

4.3 Format

The format in which the test image is to be stored in the archival medium should be straight forward resembling typical digital data so that no hardware/software modifications are required to the recording and reading devices. The following basic guidelines have in the past provided excellent image data transfer.

a) Each picture element is stored as an 8-bit data word with the transport adding, checking, and removing a parity bit in the process.

b) Each picture line is stored as a record with the transport adding, checking, and removing a record parity code.

c) Luminance and chrominance should be stored as separate segments rather than being intermixed. The luminance image should be stored first followed by the R-Y and then the B-Y image. This procedure simplifies any maintenance and diagnostic processes which may be required.

4.4 Selected Recording Medium

The conclusion reached in the previous discussion is that 8-track digital tape is the best presently available archival storage medium for retaining the standardized test images. It features low cost, high transportability, compatibility with standard transports, precise reproducibility, and long life. In addition, it lends itself well to the recommended storage format.

5.0 Playback Medium

5.1 Purpose

The purpose of the playback medium is to provide a video signal source of test images for the test and evaluation of still frame video transmission systems. The signals derived from the playback medium will be analog versions of the test images stored in digital form on the master archival tapes. The analog test image signals will be required quite often during the design and test of the video transmission systems so that they will probably degrade over a period of time due to aging and wear. New copies of the playback medium can then be generated from the digital master archival tapes providing consistent signal quality to facilitate test and evaluation.

Ideally the playback medium should be the highest quality available. This is necessary when the purpose of the test image signal is critical evaluation of a particular system or a comparative evaluation of two or more systems. In these cases the highest quality signals which are as free as possible from artifacts introduced by the storage medium are required. Highest quality is generally accompanied by highest cost. For the general test of a still frame video transmission system this cost may not be warranted and test signals which are highly controlled and repeatable but may fall somewhat short of the highest quality in all categories (such as S/N) may be totally acceptable.

Therefore two basic types of playback media are suggested:

a) very high quality limited to playback on a specific type machine to assure identical quality signals for critical and comparative evaluation,

b) high quality for playback on the vendor's designated machine to provide consistent quality signals for test and development.

5.2 Types

Two general types of playback media are possible candidates for playback of the test images; video tape and video disc. From a practical standpoint the flexibility of video tape and the widespread availability of video tape transports makes magnetic tape a very attractive medium as compared to video discs.

Video tape systems can be classified in many ways: one is by tape size and recording method. The highest quality video systems use a large (2 inch wide) tape and scan the tape crosswise using four recording heads. This is known as a quadraplex machine. The quality is excellent, but the transports are expensive and complex. Newer tape machines using 1 inch tape and helical one field per head scanning system and recording in a Type C format will provide essentially the same quality signal. They are less complex and less expensive. Both of the preceding types of video tape machines are used in critical broadcast application. Somewhat lower in quality are the 3/4 inch video cassette machines. Machines of this type are often used for broadcasting new clips, etc. The quality of signals produced by these machines is adequate for development and test purposes. The machines are also substantially lower in cost. The 1/2 inch video tape machines can provide signals with quality adequate for many purposes but will not be considered here.

5.3 Format

The format of the video signal which is the input to the video transmission system to be tested or evaluated will in almost all cases be an NTSC type signal. This basic assumption is made because the television camera used in applications employing video transmission systems will almost universally provide NTSC output signals, at least in the United States. For other areas of the world PAL or SECAM format may be more appropriate.

The basic assumption that an NTSC signal is the expected input format for the video transmission system also implies that any signal processing required be provided as part of the codec system. (For example decoding to Y, R-Y, B-Y).

5.4 Selected Playback Medium

The playback medium selected for the test image is video tape. For critical and comparative tests the 1 inch type C video tape transports are suggested. They are capable of providing broadcast quality NTSC type video signals. They are free from artifacts which may give false indications as to the performance of video transmission being evaluated.

For general use in development and test of video transmission systems the video test images should be made available in the form requested by the vendor such as 3/4 inch video cassettes.

6.0 Still Frame Image Upgrade

It is essential that, at this stage in digital still frame image transmission, test image selection should be kept flexible. The images have been selected with great care to be certain that they will provide the information necessary to evaluate the capabilities of an equipment for the transmission of still frame color images. However, newer technology, coding techniques, types of images to be transmitted, higher resolution, etc., may well require the upgrading of the test images selected.

The two-step method of generating the test tape is very flexible and will permit producing a wide range of new images photographically, live, or electronically. Therefore the system is flexible and will adapt to additional requirements.

7.0 Test Images for Motion Codec Evaluation

7.1 Purpose

The purpose of the test images for motion is to provide a "standard" source of test signals for the evaluation of the performance of codecs designed for transmission of scenes containing information in motion such as encountered in teleconferencing. The word "motion" is used here in the broad context encountered in imagery transmission. This context is perhaps best defined by interpreting the word "motion" as "change in image content." In this sense, motion encompasses any change to the information content of an image. Thus motion includes a wide range of occurrences.

A person walking across the field of view in the originating set is obviously motion. However, changes of individual data characters in an alpha-numeric display are, in this sense, also motion. The following list will provide an indication of the range of the meaning of "motion":

a) Variation in the chrominance or luminance of the scene background or foreground.

b) Changes in chrominance or luminance of alpha-numeric symbols.

c) Changes in apparent size of objects in the scene due to zooming, etc.

When area, rate of change, type of change, etc., are considered, it is apparent that motion images constitute an essentially infinite set and so the need for standard motion sequences is obvious.

The various coding techniques used in motion codecs typically perform differently with different types of motion; each having its own advantages and disadvantages. Therefore it is essential to test codecs universally with a standard set of test images which are meaningful to the evaluation in that they represent the type or effect of motion encountered in teleconferencing. Furthermore, if the test signals were being generated by different types of video sensors, they would differ markedly. This is in addition to the static differences described in the consideration of still frame test images. Assuming that the scanning standards are fixed, the motion-capturing capability of a TV camera is a function of the retention of the camera tube. Retention varies with the type of camera tube used. It may vary further with the specific potentials applied to the various control electrodes within the tube, by the light level applied to the target, and, to some degree, the temperature of the target. Therefore it is essential that there be a standard test signal source for the evaluation of motion codecs.

7.2 Parameters to be Tested

The parameters to be tested include all of those listed in Section 2 for the evaluation of still frame codecs. These parameters are fundamental. The recommended method of evaluating the performance of the codec in transmitting these parameters is by means of the still frame test images and techniques previously described.

The evaluation of motion performance requires an additional set of test image sequences. In the motion case, the image consists of two parts; that which is in motion (changing) and that which is static. The ratio of the relative size of these two parts as well as the rate of change incurred by the motion can greatly affect the equipment performance. All of these factors must be considered in the selection of the test image sequences.

Basically then, considering all the previous factors, the amount of motion can be defined as the number of picture elements and the degree to which they change from one field to the next or within 1/60 second. It is essential to keep this concept in mind although it is difficult to apply to a practical evaluation.

7.2.1 Static Parameters

The static parameters to be tested are tabulated in Section 2. Test images for their evaluation are also defined in Section 2. Test images for their evaluation are also defined in Section 2. These parameters are so basic that they apply just as well to motion codecs as to still frame codecs. Therefore all of Section 2 applies to the evaluation of motion codecs.

7.2.2 Resolution, Gray Scale and Color Rendition of the Static Area of a Scene Containing Motion.

Many factors, internal and external to the codec, affect the quality of the motion area of a scene being transmitted by a codec. These will be described in Section 7.2.3. In a practical teleconferencing situation, the static area may be of prime interest even while there is substantial change in the overall scene content. The quality of the static area of the image can be markedly affected as a function of image change, depending on the coding techniques incorporated into the motion codec. Therefore the test image sequence must contain a static area as well as an area with motion. Figure 7-1 is a chart designed for this purpose. The static area contains resolution bars ranging from 100 to 500 TV lines referred to the picture height. These bars will provide quantitative information regarding the ability of the codec to reproduce static information while the remainder of the image changes at various controlled rates. The static area also contains a color bar segment so that both gray scale and color condition can be thoroughly evaluated during controlled motion sequences. The test motion sequence proposed contains an interval during which there is no motion so that the evaluators can establish a qualitative and quantitative reference of the static area performance. This is followed by an interval during which the background changes at a predetermined rate. Several segments of sequences of this type at various motion rates will permit thorough evaluation of the static performance of a motion codec while transmitting motion scenes.

7.2.3 Performance of the Motion Area of a Scene

The same test image sequence described in Section 7.2.2 and shown in Figure 7-1 can be used to evaluate the performance of the motion area

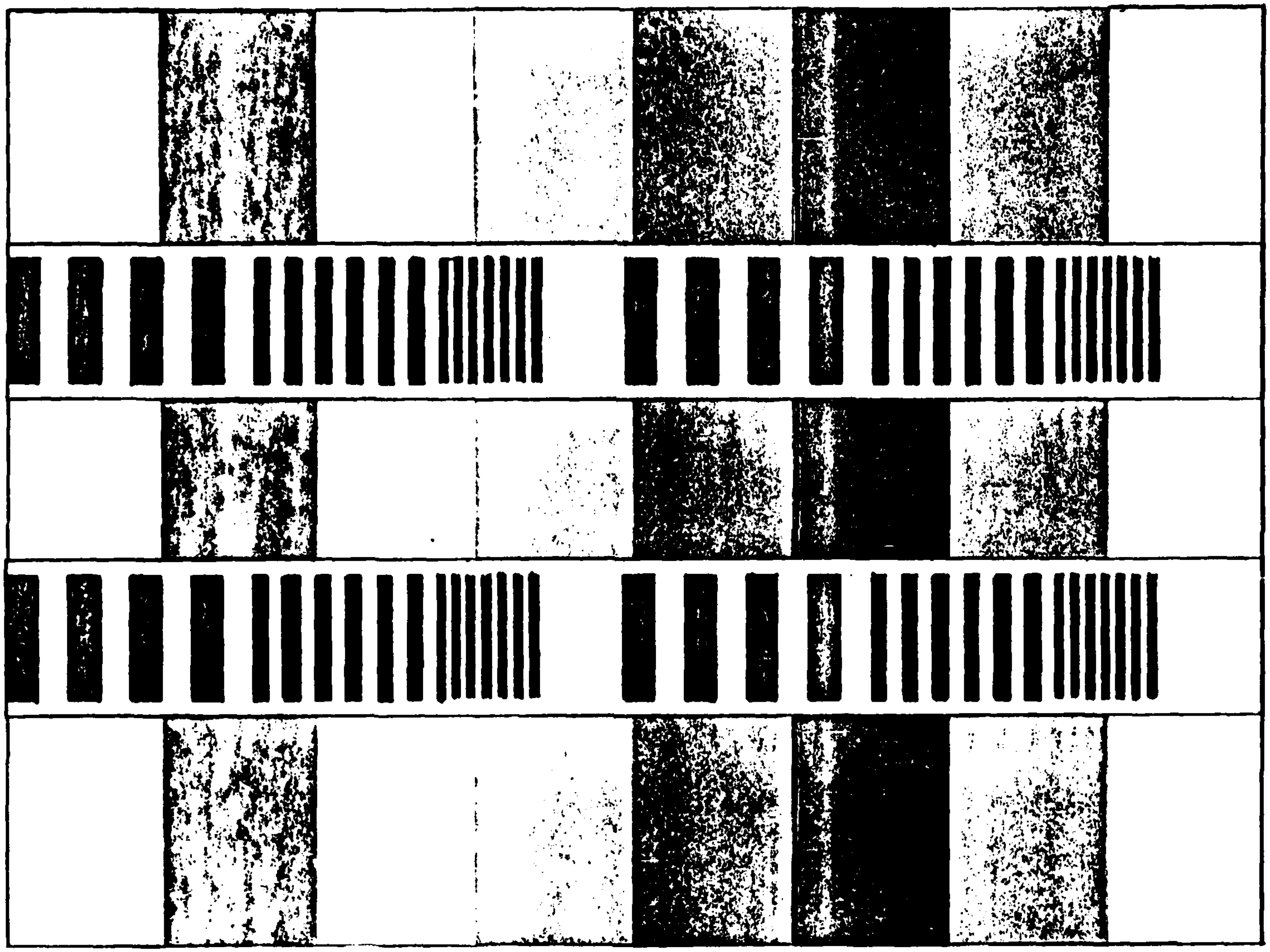


FIGURE 7-1 COLOR MOTION TEST IMAGE

of a scene particularly with regard to resolution, gray scale, and color rendition.

Generating these sequences by means of a television camera would at first seem to be the proper method to be employed because in an actual teleconferencing situation a television camera is the most likely input sensor. However, television cameras substantially affect a system's capability to reproduce motion scenes particularly when high scene content change rates are involved. The test sequence to be described utilizes an electronically generated background which is free of the motion-degrading effect of television cameras. At the time that these sequences are actually generated, a re-evaluation to the relative merits of these two approaches should be made based on experimental data.

The test image shown in Figure 7-1 consists of four parts as listed below:

<u>Background</u> (Motion)

Resolution Bars
Color Bars

<u>Foreground</u> (Static)

Resolution Bars
Color Bars

The static segment, as previously described, comprises a rectangular segment of the image area which is 25 percent of the picture in height and full width.

It contains resolution bars ranging from 100 to 500 lines identical to the resolution strip in the motion section. Immediately below the resolution bars within the static segment is a complete set of color bars in the standard configuration: white, yellow, cyan, green, magenta, blue, and black. When the test sequence shows no motion the static color bars are aligned with identical color bars in the motion portion of the image.

The motion portion of the image also consists of resolution bars and color bars identical to those in the static segment.

The motion segment occupies the lowest 20 percent of the image area. Above this is the static segment which occupies the next 25 percent. The upper 50 percent of the image consists of motion segment. Part of this upper segment of motion is a band of resolution bars, 15 percent of the image in height and full width.

When the sequence contains no motion, the color bars as well as the resolution bars in the two segments are aligned vertically. When the sequence contains motion, the motion segment appears to move to the left. Geometrically this consists of the lowest 20 and uppermost 55 percent for a total of 75 percent of the image area.

Since the motion area consists of eight equal-width, vertical strips containing totally different information, a lateral shift of 1/8 of the picture width will cause all of the picture elements in the motion segment of the image to change. This represents a change in 75 percent of the total picture element. If this change occurs every field, the motion can be considered continuous. It is anticipated that a continuous change of 75 percent of the area of an image is highly unlikely in a teleconferencing application and therefore is beyond the necessary upper bound. If the image shifts to 1/16 of the picture width, 37.5 percent of the image will change. The following is a tabulation of image motion per field and percent area change per field anticipated as useful for motion codec testing.

Shift (% of Pix Width per field)	Motion (% of Area Change per Field)
1/8	75
1/11	37.5
1/32	18.75
1/64	9.38
1/128	4.69
1/256	2.35
1/512	1.18

The use of the bars distributes the motion to vertical strips spaced throughout the image area.

The magnitude of the picture change can be further defined as to the range over which each pel changes. For example, if a picture element changes from black to white, a 100 percent dynamic range change has occurred. If a pel changes by only 10 percent in gray level, the change is less substantial. Depending on the codec architecture, changing the stored and/or transmitted digital value of a pel may be affected by the magnitude of the change in the pel. In a typical image, 100 percent magnitude changes within a field interval are not considered common. However, they may occur if camera retention is omitted. A typical occurrence is in the operation of a video switcher when punching from one scene to the next. In addition to the luminance or gray scale changes, changes in chrominance must be considered.

The image designed for this test provides "average" changes. Luminance ranges from white to black through the luminance values of the six intervening color bars. Chrominance varies through 360 degrees with jumps from and to zero chroma at either end. In addition, the resolution bars show the effect of the codec on 100 percent luminance changes for small areas.

7.2.4 Qualitative Motion Evaluation

The preceding section described a method of evaluating the effects of motion on the quality of the reconstructed image produced by a motion codec. It used image sequences containing highly defined and controlled amounts of motion. This technique can be used to give the codec a semi-quantitative evaluation factor. It uses a totally artificial scene. This section will describe motion sequences which are typical in a teleconferencing application. These can be used to generate a meaningful qualitative evaluation factor.

Occurrences in televising teleconferencing scenes which produce motion or, more precisely, changes in image content include the following test sequences for their evaluation. In addition, one new image (not shown) is suggested to provide a highly articulated scene of saturated colors. Sharp demarcations in chrominance and luminance are provided. The scene is a set of flags of various motions selected and illuminated to produce the features described above.

7.2.4.1 Punch

This is a video switcher operation which selects the video signal from one television camera to replace that from another television camera for transmission. The switching operation occurs during the vertical blanking interval so that the picture is completely changed from one frame to the next. This is the "most violent" change or motion which can occur in the system. A punch sequence changing from a black scene in one frame to a white scene in the next, for example, requires all luminance pels to change full magnitude in 1/30 second.

The following sequences consisting of punching back and forth between the pictures listed are recommended.

Picture 1	Figure	Picture 2	Figure
Single Person	2-4	Two People	2-5
Single Person	2-4	Color Bars	2-14
Black Field		Single Person	2-4
Black Field		Color Bars	2-14
Two People	2-5	Bar Graph	2-12
Single Flag		Full Set of Flags	

7.2.4.2 Wipe

A wipe is a change in scene content produced by the special effects position of a video switcher. The change in scene content is effected by replacing the video from one camera with the video from another camera on a geometrical basis (referred to the displayed picture). Thus segments of the first picture are replaced by those of a second picture as the wipe progresses. The simplest wipe is a lateral wipe in which the first image is replaced from the left as the second image starts to expand from the left in a sort of "windshield wiper" manner. The images themselves may be static: the "motion" is provided by the change in picture elements which occur during a wipe. This is a far less "violent change" than a punch.

The following sequences of wipes between the pictures listed at the rates listed are recommended. The form of the wipe is not very important but a wipe starting with a small square at the center of the picture and growing outward until it replaces the original scene is very common and is suggested.

Picture 1	Figure	Picture 2	Figure 2	Wiper Rate (Seconds)
Single Person	2-4	Bar Graph	2-12	0.5/1.0/2.0
Single Person	2-4	Color Bars	2-14	0.5/1.0/2.0
Two People	2-5	Single Person	2-4	0.5/1.0/2.0
Black Field		Color Bars	2-14	0.5/1.0/2.0
Black Field		Single Person	2-4	0.5/1.0/2.0

7.2.4.3 Zoom

Zoom is the change in the image magnification produced through the lens of a TV camera or through an electronic manipulation of the scheme used to read an image out of a frame-store memory. The motion in this case is the change induced in picture elements as the image appears to grow or shrink. As an example, consider zooming in on a small black square set on a white background. As the change in magnification causes the black square to grow until it fills the total display area, all of the originally white picture elements must change to black.

The following zoom sequences are recommended.

Picture	Figure	Rate (Seconds)
Single Person close-up to full image	2-4	0.5/1.0/2.0
Two People close-up to full image	2-5	0.5/1.0/2.0
Single Flag to full set		0.5/1.0/2.0

7.2.4.4 Mix

Mix is another method of selecting the video being transmitted from one camera to another by means of a video switcher. In this technique the original signal is gradually decreased in magnitude as the second video signal is increased and added to the first signal so that the overall amplitude remains constant. The visual effect is that of one picture fading away and being replaced by a second which fades in to full. The following are recommended mix sequences.

Picture 1	Figure	Picture 2	Figure	Rate (Seconds)
Single Person	2-4	Graph	2-12	0.5/1.0/2.0
Two People	2-5	Graph	2-12	0.5/1.0/2.0
Single Flag		Full set of Flags		0.5/1.0/2.0
Single Person	2-4	Two People	2-5	0.5/1.0/2.0
Single Person	2-4	Color Bars	2-14	0.5/1.0/2.0

7.2.4.5 Conventional Motion

The sequences previously described provide controlled motion which will occur in teleconferencing systems because of the mechanics of the actual telecast production. In addition, movement of objects in the scene must be considered as a codec evaluation tool.

First, it is recommended that a scene very high in both luminance and chrominance contrast and continuing motion be included. The flag scene previously described is excellent. The flags mounted on a turntable will provide motion at various rates within a single scene as the set revolves.

Second, it is recommended that scenes with people in motion at fairly slow rates be included. Highly controlled and carefully produced motion picture films of the sets and people used in the still-frame test images (Figures 2-3 through 2-6) are available. These, or new sequences, patterned after these films and produced to the same standards are recommended.

7.3 Motion Quality Rating

The rating system used for still-frame codec evaluation will satisfy the motion quality evaluation requirement. However, there are additional parameters to be evaluated beyond those which are identical to the still-frame case.

The following are included in this category.

* Image breakup in motion/stationary area.
* Image smear in motion.
* Resolution in motion/stationary area.
* Luminance performance in motion/stationary area.
* Chrominance performance in motion/stationary area.

8.0 Motion Digital Coding Recommendation

Based on many years of experience with digital video systems, it is recommended that the digital coding parameters previously defined will be ideal for the coding of the motion test sequences. This has been verified by the SMPTE and the EBU in their selection of digital coding standards.

The reader is, therefore, referred to Section 3.3.

9.0 Performance Evaluation Criteria/Procedures

9.1 Parameters to be Evaluated

The object of the performance evaluation is to determine the performance of the image transmission system in the transmission, reconstruction, and display of pictorial data typical of teleconferencing applications. This evaluation can be made with the use of a few well-designed test images. The images are selected so that the transmission of their content will thoroughly stress the various technical parameters of the equipment. Certain of the test images, such as test charts, evaluate certain parameters quantitatively; eg, resolution. Others provide pictorial data which permits qualitatively evaluation of the same or other parameters; eg, facial recognition or flesh tones.

The technical parameters to be evaluated directly or qualitatively are tabulated below.

<u>Quantitative</u>

1. Gray Scale
2. Resolution
   - 2.1 Horizontal
   - 2.2 Vertical
   - 2.3 Slant Edges
3. Quantization
4. Color
   - 4.1 Saturation
   - 4.2 Purity
   - 4.3 Hue Accuracy

4.4 Crosstalk

4.5 Effects of Color on Luminance

4.6 Effects of Luminance on Color

5. Signal-to-Noise Ratio

6. Temporal Effects

6.1 Frame Grab

6.2 Retention of Preceding Image

Qualitative

1 Flesh Tones

2 Recognizability of Subjects

3 Effects of Scene on Appearance of Subject

4 Clarity of Print, Graphics, and Maps

5 Ability to Specify Single Color

6 Ability to Determine Color in Multi-Color Scenes

7 S/N or Artifact Performance

The evaluation of these parameters will give a clear indication of the capability of the equipment being evaluated.

9.2 Instructions on the Use of Tape as the Input Source

Evaluation of the equipment performance will always be on two bases.

Absolute Basis. This technique evaluates the system output as compared to its input. It determines how adept the system is at transmitting the particular input image being used. This, of course, is important, but it is often of equal concern to determine the equipment performance as compared to its own previous performance or the performance of other equipments tested at other times and locations. To do this it is essential to use a standardized set of test imagery which is always consistent in quality. This can be achieved with the test tape approach.

Comparison Basis. This test compares the displays of two equipments to determine their relative capabilities. It is essential to compare two equipments using the same input signals and that these be relevant to teleconferencing. Since it is often necessary to repeat tests or segments of tests, the test signals must be exactly reproducible. This can be achieved with a standardized set of test imagery on a test tape.

Finally, the test tape provides consistent test imagery so that all properly conducted tests performed using it will provide a means of cross-correlating the performance of all equipments tested.

The evaluation tests should include the following precaution. First, since a single tape and transport is used to provide the input signal the quality will be consistent. It should be verified that the signal as it is applied to the equipment or equipments is of the proper amplitude and equal quality. Second, the test area should be kept free from distractions caused by operating equipment and unnecessary personnel. It is desireable that only the display monitor (s), before/after or A/B switch, together with comfortable seating and writing facilities be located in the evaluation area. Third, the viewing should be arranged so that the monitor is observed from about 5 times picture height. CCIR Recommendation 500 is suggested as a guideline. A gray surround is suggested in the viewing area with the luminance behind the picture monitor less than 15% of the peak picture luminance. It is important to prevent any lighting provided from falling directly on the monitor screens and in particular from causing a specular reflection to the viewer. In general, the black to peak white ratio should be less than 0.02. The lighting should be diffuse and adjusted to match the output and color of the monitor when displaying mid-gray. This eliminates the need for constant eye adaptation as the viewers look away and back to the monitors. The monitors should be at eye level so the viewer looks squarely at the screen. The monitors (if not an integral part of the equipment) must be of adequate quality so that they will not degrade the output signal of the equipment under evaluation. A light output of 50 foot-Lamberts is recommended.

9.3 Methods of Evaluation

9.3.1 Evaluation of a Single Codec

Single codecs may be thoroughly evaluated on the basis of an input/output comparison. The video signal of the standard test images directly from the test tape is displayed on the monitor.

The output signal of the codec (digitized, processed, stored, and converted back to analog) is then displayed on the monitor. The evaluator(s) then rate the comparative quality of the output image relative to the input image.

9.3.1.1 Interconnect

The equipments are interconnected as shown in Figure 9-1. The video tape recorder provides a composite output signal of the test images to a distribution amplifier. The distribution amplifier provides two identical otuput video signals of 1.0 volts peak to peak amplitude. The quality and amplitude of these signals must be verified by means of an oscilloscope. One of the video signals is connected to the terminated input of the codec under test. The other video signal is connected to one input of a video switch (effectively a single pole double throw configuration). The output of the codec is connected to the other input of the video switch. The output of the switch feeds the display monitor. By means of the switch either the input or the output of the codec can be viewed on the monitor for evaluation. The switch must be configured to terminate the unused input and to minimize crosstalk.

9.3.1.2 Procedure

It is recommended that the population of evaluators contain both experts and non-experts. Experts tend to be very critical while non-experts represent the majority of users and therefore tend to evaluate the system performance with its eventual use in mind.

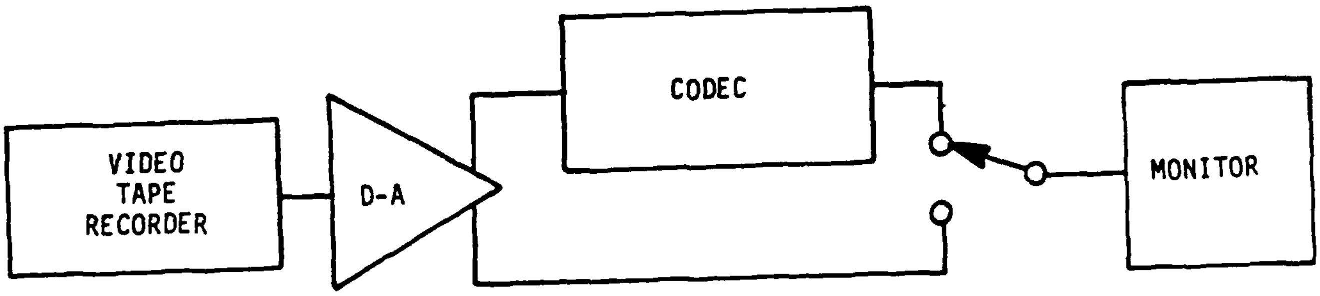


FIGURE 9-1 SINGLE CODEC EVALUATION

Both are necessary. It is further recommended that test charts not be evaluated by the non-experts but be thoroughly evaluated by experts who can properly interpret the monitor display.

The number of evaluators depends on the confidence level needed in the results of the test. For a thorough evaluation of the equipment performance 10 experts and 20 non-expert evaluators or more are recommended.

The test procedure is relatively straightforward. It is essential to brief the evaluators before the test begins as to what is expected of them. It is not necessary that the evaluators be highly technically qualified. It is well to include user type personnel concerned with qualitative performance in addition to those adept at evaluating quantitative performance. Of particular importance is giving the evaluators some experience in rating performance using the comparison rating scale.

The procedure is to alternately display the input and the output video signals on the monitor so that the evaluators can compare the two displays and assess a comparative rating to the output signal. The period of evaluation continues until the evaluators have each independently reached a decision. This process continues through all of the test images selected for evaluation.

The ratings of the individual evaluators are then tabulated for a final conclusion.

The development of a single rating factor is still under consideration. However each of the individual parameters as previously described can be rated as the average of the individual ratings by each of the evaluators for that parameter.

9.3.1.3 Rating Scale for Single Codec Evaluation

As discussed in the previous sections, the tests are of two types; quantitative and qualitative. The quantitative tests permit the evaluator to determine an absolute numeric value to define the system performance; for example, resolution. It may be well to state also the difference noted between display of the input and output signals.

The qualitative evaluation can be stated in terms of the apparent impairment of the output display as compared to the input display. The five-point impairement scale as defined in CCIR recommendation 500-1 is proposed as follows.

5	Imperceptible	(Excellent)
4	Perceptible, but not annoying	(Good)
3	Slightly annoying	(Fair)
2	Annoying	(Poor)
1	Very Annoying	(Bad)

It is very important to thoroughly instruct the evaluators in the use of this scale.

9.3.2 Comparison of Two Codecs

It may be necessary on occasion to compare the performance of two codecs. The evaluation can be performed by one of two techniques: comparing one codec to another, or evaluating each independently as in Paragraph 9.3.1 and then comparing the results.

Comparison of one codec to another can also be rated by means of the following scale.

+3 Much better

+2 Better

+1 Slightly better

0 The same

-1 Slightly worse

-2 Worse

-3 Much worse

9.3.2.1 Interconnect

Three interconnects are described;

a) two terminals with switch

b) two terminals with two monitors, and

c) two terminals and an input reference

The video signal from the test image tape is applied to a distribution amplifier which provides two or three output signals which can be adjusted to be identical. See Figure 9-2.

In the first, (A), arrangement, two codecs are fed with identical signals of the test images. The output signals from the codecs are applied to the single pole-double throw switch which in turn feeds a single monitor. Evaluation is performed by viewing the signals from each of the two codecs alternately for comparison.

In the second, (B), arrangement the switch is eliminated and

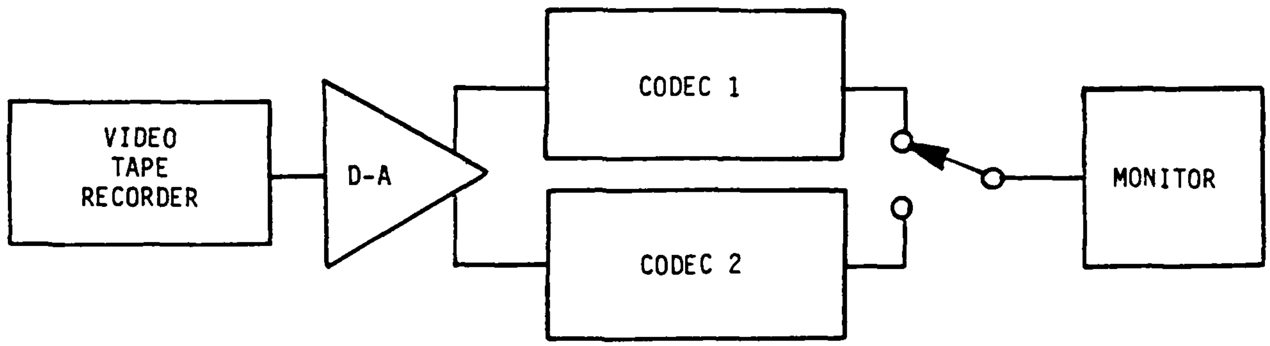


FIGURE 9-2A COMPARISON EVALUATION SET-UP

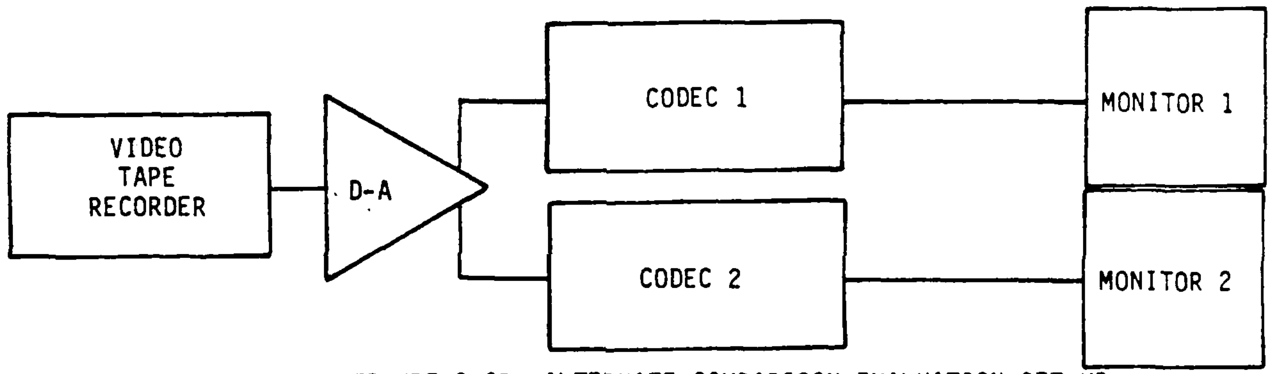


FIGURE 9-2B ALTERNATE COMPARISON EVALUATION SET-UP

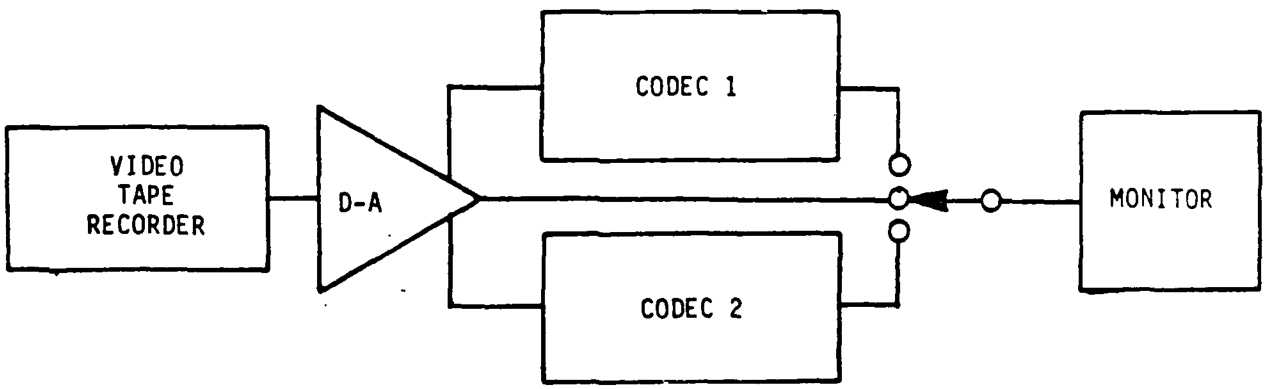


FIGURE 9-2C RECOMMENDED COMPARISON EVALUATION SET-UP two monitors are used each displaying one of the video signals from the codecs. Great care must be taken to be sure that the two monitors are identically and properly adjusted.

In the third (C), and preferred, arrangement a single pole-triple throw switch is used to select either one of the two codec output video signals for display on the monitor. Evaluation is performed by rating each of the two monitors against the input signal.

9.3.2.2 Procedure

The proceaure for the A and B arrangements are similar to that used in the evaluation of a single codec. The evaluators compare the two monitor displays and rate one with respect to the other. The rating scale shown below may be used. The better of two images becomes the temporary standard against which the other is rated.

The preferred procedure, C, is essentially identical to that used in the evaluation of single codec. Both of the codecs are evaluated against the input standard. The results for each parameter are tabulated. The rating factors for each codec are summed, divided by the number of entries to determine an average rating factor. The result is an absolute rating. The difference between the average rating factors is the relative rating between the two codecs.

9.3.2.3 Rating Factor

The rating factors for the A and B evaluation techniques are shown below.

5 Difference is Imperceptible

4 Difference is Perceptible, but not annoying

3 Difference is slightly annoying

2 Difference is annoying

1 Difference is very annoying

The rating factors for the C evaluation technique is the same as that for a single codec as shown below.

5 Imperceptible

4 Perceptible, but not annoying

3 Slightly annoying

2 Annoying

1 Very Annoying

10. Recommended Method of Generating Archival and User Test Tapes

The recommended method of generating the digital archival tape and the user test tape is a precise though very straight forward technique. The process is illustrated in Figure 10-1. The test images (opaque or transparencies) are viewed by a properly aligned TV camera producing an NTSC format output video signal. The illumination of the test image material and the adjustment of the TV camera are key items in the process. Since the system is essentially digital subsequent to the generation of the video signal, the original signal quality will be retained. The video signals need be generated only once since additional identical copies of the digital tape can be produced without change in quality. Therefore considerable effort can be expended in generating an optimum initial video signal. This portion of the process is conventional and hardware is readily available.

The second phase of the process is equally important but requires comparatively unconventional hardware. This phase consists of decoding the analog NTSC input video signal into its three component parts; Y, R-Y, B-Y as previously described. The signal components are then each filtered to remove any frequency component which might produce aliasing in the final product and are converted to digital form. This requires timing circuits to control the sampling rate and phase of the three signal components. The samples are then converted into digital form by three A/D converters. The Y A/D converter provides

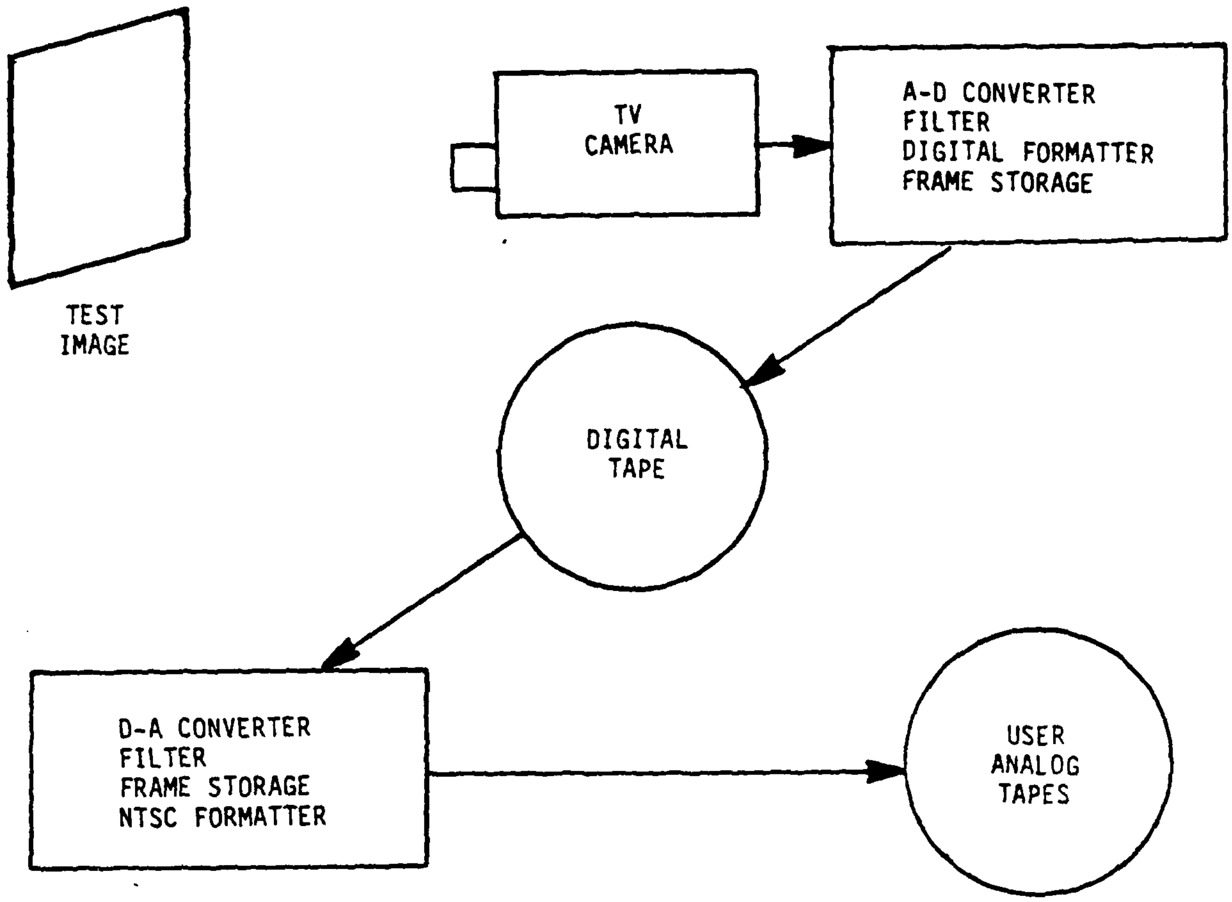


FIGURE 10-1 RECOMMENDED METHOD OF GENERATING ARCHIVAL DIGITAL TAPES AND USER ANALOG TEST TAPES

digital codes ranging from 0 to 255. The R-Y and B-Y A/D converters each provide digital codes between -128 and +128. Thus all codes have a common length of 8 bits. The three sets of digital data corresponding to a single video frame are stored in a digital frame memory. The frame memory performs two functions. First, it organizes the intermixed Y, R-Y, and B-Y digital video data so that it can be transferred to the digital tape in three segments, first Y, then R-Y, and then B-Y. Secondly it performs the speed buffering from the real time video data rate to the data rate accepted by the digital tape transport.

One device which performs that function as well as the inverse function to be described subsequently is available at Delta Information Systems. It was designed by members of the technical staff and has proven itself for about 5 years as a digital video transmission system. Its basic parameters are described as Standard #2 in Section 2. This device is available for the generation of the digital archival tapes and the subsequent generation of the analog user tapes.

The process of recording the digital tape can be performed directly as shown with some minor modifications. However by means of the simple expedient of transferring the data from the converter to the tape through a computer with a PR-11B interface card the system can be used as it was originally designed and as it has been previously used. This interface is also available.

The inverse process of converting hte digital video tape into an analog NTSC video signal is equally straight forward with the converter described. The digital data is transferred from the digital tape to the frame memory in the frame storage of the converter at data rates compatible with the tape transport. The converter D/A converts the Y, R-Y, B-Y components of the video signal into analog form simulataneously at real time video data rates. Each component is properly filtered and the three are then encoded into a real time NTSC signal. This is done in a continuous manner to generate a continuous NTSC video signal at real time rate. Each still frame can therefore be recorded on the user analog tape in segments of any length desired.

Since the analog signal is a true continuous NTSC signal it can be recorded on any compatible video tape recorder. For test tapes that can be a 3/4 inch video cassette recorder or for critical application a 1 inch broadcast quality type "C" format video tape recorder.

Other hardware systems which can perform this function may be available but their performance must be thoroughly investigated before they could be recommended for the generation of the standardized test tapes.

11. Recommended Efforts in the Development of Standard Imagery for Teleconferencing

11.1 Discussion

During the course of this study, a preliminary investigation of the use of higher resolution than 525 line television was conducted. There is considerable activity by several vendors and standards organizations to develop high resolution television techniques. Many of these techniques are grouped under the name of High Definition Television (HDTV). The Society of Motion Pictures and Television Engineers (SMPTE) has several committees investigating HDTV with a goal of trying to establish standards before a proliferation of different, non-compatible equipments is established. The following proposed efforts describe various tasks which should be undertaken in order to be prepared for the use of HDTV in teleconferencing applications and to initiate the establishment of recommendations for standards for the Federal Government.

11.2 HDTV Survey and Analysis

It is recommended that a survey of the literature be performed on HDTV techniques. Also, vendors making HDTV equipments should be contacted to gain technical information on their products. This information will serve as a technical base to supplement information available from the various SMPTE committees on HDTV. From these data, an assessment of the HDTV technology will be made to determine its impact upon teleconferencing applications.

11.3 Uses of HDTV in Teleconferencing Systems

A study should be made to define the total HDTV Teleconferencing System. The following system components should be considered in the study:

- Cameras
- Monitors
- Large Screen Displays
- Codecs
- Transmission Rates
- Data Links
- Recording Techniques and Equipments

The result of the study will be the generation of system diagrams and system parameters from which areas can be identified for future standards activities.

11.4 Coordination with Standards Organizations

As previously mentioned, there is considerable activity in HDTV technology in various standards organizations. Because of this rapid development of HDTV technology and its applications to teleconferencing, it is recommended that active coordination and participation be maintained with the key standards groups including the following organizations:

- SMPTE
- CCIR

CCITT

IEEE

EIA

This participation should expedite the timely disemination of proposed HDTV standards which could bear upon the effort to establish U.S. Federal Standards for digital video teleconferencing.

N